

シラバスの文書クラスタリングに基づく カリキュラム分析システムの構築

野澤 孝之[†] 井田 正明[†] 芳 鐘 冬 樹[†]
宮崎 和光[†] 喜 多 一^{††}

高等教育機関が独創的なカリキュラムを設計しようとする場合や、第三者が高等教育機関のカリキュラムの特徴を評価する場合、多数の教育機関にまたがる教育内容の横断的な把握が必要となる。しかしこれは専門家にとっても負荷の高い課題であり、カリキュラム設計や評価の方針を立てやすくするためのコンピュータを用いた支援環境が望まれる。本研究では、共通形式化されたシラバスデータを対象に、それらが含む専門用語を抽出し、その出現頻度に基づき科目間の類似度を計算しクラスタリングを行い、多角的な分類軸に沿って科目のクラスタへの帰属分布を視覚化し対照することで、カリキュラムの特徴把握を支援するシステムを開発した。

Construction of Curriculum Analyzing System Based on Document Clustering of Syllabus Data

TAKAYUKI NOZAWA,[†] MASAOKI IDA,[†] FUYUKI YOSHIKANE,[†]
KAZUTERU MIYAZAKI[†] and HAJIME KITA^{††}

In designing of an original *curriculum* by a higher education institution, or in external evaluation of an institution's curriculum, comprehending the curriculum contents of many institutions in the same field is necessary. However, this is very hard to do even for education experts in that field. Therefore, some computer system which promotes discovering the features of the curriculums is desired by the curriculum designers and evaluators. In this research, we developed a new system which supports the feature discovering of curriculums, utilizing the *document clustering* technique. This system treats *syllabus* data which constitute some curriculums and are expressed in a common format, calculates the similarity between the syllabi based on the occurrence frequency of technical terms, clusters the syllabi, and thus helps to find distinguishing features of the curriculums by visualizing and comparing the assignments of the syllabi to the clusters along various classification axes.

1. はじめに

現在、日本における大学をはじめとする高等教育機関をとりまく状況は、大きく変わりつつある。2004年4月から国立大学は法人化され、同じく2004年度から全国の大学等を対象にした認証評価制度が導入された。また文部科学省の「特色ある大学教育支援プログラム」は継続的に実施されているし、日本技術者教育認定機構（JABEE）による技術者教育プログラムの認定は受審への動きが活発化している。以上のような動きの背後にあるのは、知識社会時代を迎えて高等

教育機関に対する期待が増大していることと、それゆえの備えるべき水準を満たしたうえでの個性・特徴の進展という要望であろう。

この状況の中で、カリキュラム（教育課程）の設計・分析という作業が持つ重要性も増してきている。高等教育機関は、共通的な一定要件を満たしたうえで、学問や技術の進歩・社会ニーズの変化に合わせながら（たとえば研究1）を参照）、学生にとってより魅力ある独創的なカリキュラムを設計し、改善し続けることを求められる。また高等教育の質について外部評価を行う第三者評価機関の側からも、各教育機関が持つ目的目標に則した評価という観点から、カリキュラムがどのような特徴を持つかを分析し、それが教育機関の目的

[†] 独立行政法人大学評価・学位授与機構
National Institution for Academic Degrees and University Evaluation

^{††} 京都大学
Kyoto University

大学は教育だけでなく研究も使命とする機関であるが、本研究では教育機関としての側面を抜き出して扱う。

目標の方向に合致するものであるかを判断することが必要となる。もちろん教育サービスの顧客となる学生にとっても、各教育機関が提供するカリキュラムの特徴を明らかにするデータはサービスを選択するうえで有効な判断材料となると期待される。

独創的なカリキュラムを設計しようとする場合や、カリキュラムの特徴を評価する場合、多数の教育機関にまたがるカリキュラムの横断的な把握が必要である。そして実際に体験できるカリキュラムに限りがある中で、横断的な把握を得るための有力なデータが、カリキュラムを構成する科目の内容を記述したシラバスである。しかし複数機関のカリキュラムを構成する科目のシラバスを集めると膨大な量となり、またそれらは毎年といった頻度で改訂されるため、それらを読み通したうえで各カリキュラムの特徴を抽出するのは非常に負荷の高い困難な課題となる。そこで著者らは、大量のシラバスデータを分析して特徴を抽出し、カリキュラムの設計や評価の方針を立てるのに役立つシステムが望まれると考えた。

本研究では、シラバスデータを対象に、それらが含む専門用語の出現頻度に基づきシラバス(科目)間の類似度を計算し、クラスタリングを行うシステムを構築した。教育機関別、必修/選択の区分、履修年次などを組み合わせた分類軸に沿ってシラバスのクラスタへの帰属分布を視覚化することで、カリキュラムを様々な角度から分析することを可能とした。

なお、類似の要素技術を用いてカリキュラムの特徴を構造化・可視化した研究として、研究2)がある。ただし研究2)が単一カリキュラム内の詳細構造の検討に重点を置いたものであり、また科目の特徴判定などに人の判断が介在するのに対して、本システムはカリキュラム単位での比較とポジショニングに重点を置いており、また科目の特徴判定を自動で行うため幅広いカリキュラムを対象にした分析が手軽に行える点が異なる。

次章ではデータ分析における本研究の位置付けを明らかにする。3章では本研究で構築したシステムによる分析の流れを説明する。4章で本システムによる分析例を示す。最後に5章でまとめと今後の課題を述べる。

2. 関連研究と本研究の位置付け

データ分析の分野の中で、個々のデータに分類構造を与えるデータ組織化の研究は、外的基準を前提しないクラスタリングと、外的基準に基づくクラシフィケーション(教師付き分類)とに大別される。両者の

違いはその用途の違いとしてとらえることができる。前者のクラスタリングは、有効な分析の切り口がまだ定まっていない状態で、それをデータ自身の内から発見したい場合に利用される。本研究が対象としたいのは、こちらの知識発見的な用途である。クラスタリングの手法については次章で述べる。他方で後者のクラシフィケーションは、分類のクラスがあらかじめ与えられている状況で、個々のデータがどのクラスに割り当てられるかを求めたいときに利用される。シラバスデータにクラシフィケーションを適用した研究の例として、大学評価・学位授与機構の学位授与事業において学位申請者が履修した科目を学位認定の要件に則して適切に分類する作業の支援を行う、科目分類システム³⁾がある。なお、クラシフィケーションでは一般には、決定木や階層的ニューラルネットワーク、Support Vector Machine など、教師付き学習の方法がよく用いられる⁴⁾。

クラスタリングのうち、特に文書データを対象としたものを文書クラスタリングと呼ぶ。文書クラスタリングは情報検索(IR)の分野で精力的に研究されてきた⁵⁾。それらの多くは、文書集合をあらかじめクラスタリングしておき、検索クエリへの適合性をクラスタ単位で判定しクエリ応答を作成することで、検索の性能(再現率や精度)を上げようという考えに基づいていた。しかしこのアプローチの計算コスト対効果が低いことが認識されるにつれて、文書クラスタリング活用の目的は、文書集合を閲覧者の興味に合わせて効率良く閲覧し、その全体的内容をとらえるという方向に移ってきている^{6),7)}。シラバス集合が全体(カリキュラム)として示す特徴をとらえるために文書クラスタリングを用いるという点で、本研究もこの同じ方向に沿っている。

クラスタリングはデータ間の類似度に基づいて個々のデータをクラスタにまとめる。文書クラスタリングにおいて文書間の類似度を定める方法は、テキスト内容に基づく方法と、文書間のリンク構造に基づく方法に大別される。前者は各文書が含む単語の頻度情報などから類似度を計算するものであり、シラバスという内容記述的な側面の強い文書を対象とするため、本研究で用いる文書クラスタリングもこちらに準じる。後者は学术论文の共引用関係⁸⁾やハイパーテキストのリンク構造⁹⁾などを材料に類似度を計算する。そのほか、テキスト内容とリンク構造の両方に基づき類似度

ただし5章で言及するように、科目間の履修順序関係などはシラバス間のリンク構造として活用が期待できる。

表 1 本システムの分析オプション一覧
Table 1 Curriculum Analyzing System options.

分析ステップ	オプションの種類	選択可能なオプション値
1. 分析対象選択	カリキュラム&シラバス選択条件 専門用語辞書	シラバス XML スキーマの定めるデータ項目を用いた任意の SQL 検索条件 EDR 専門用語辞書, 専門用語自動抽出システムで生成した用語リスト
2. シラバス定量化	用語抽出の対象データ項目 用語重み付け関数	科目名, 授業概要, 授業の目的, 授業計画, 教科書, 参考書, ... の組合せ 出現頻度 tf , TFIDF 値 $tfidf$
3. シラバス間 類似度計算	規格化操作 類似度の定義式	有/無 内積類似度, ユークリッド距離の逆数
4. クラスタリング	クラスタリング手法	階層併合的手法—最短距離法, 群平均距離法, 最長距離法—
5. 結果の観察	クラスタ帰属分布比較の分類軸	機関名, 学部, 学科, 必修/選択の区分, 授業形態, 開講学年, ... の組合せ

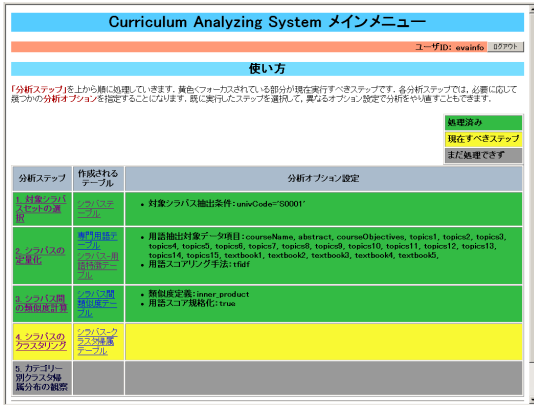


図 1 カリキュラム分析システムメインメニュー
Fig. 1 Curriculum Analyzing System main menu.

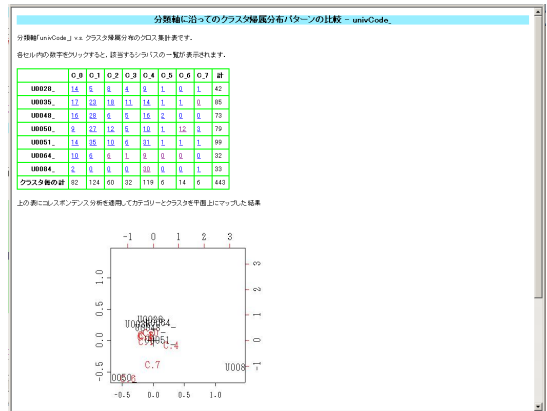


図 2 クラスタ帰属分布のクロス表とその視覚化
Fig. 2 Assignment of syllabi to clusters and its visualization.

を定める手法もある¹⁰⁾。

3. 本システムによる分析の手順

本研究で構築したシステムが提供するカリキュラム分析は、

- (1) 分析対象とするカリキュラムとシラバスを選択,
- (2) 専門用語抽出を適用するデータ項目を選択し, 含まれる専門用語に基づき各シラバス(科目)の内容を縮約・定量化,
- (3) シラバス間の類似度を計算,
- (4) 類似度に基づきクラスタリング, そして
- (5) 各シラバスのクラスタへの帰属分布を様々な軸に沿って可視化する,

という手順で行われる。(1)は分析の準備であり,(2)~(4)が文書クラスタリングの標準的な手続きに相当する。そして(5)がカリキュラム全体の特徴を観察するステップである。また,この各ステップで分析の詳細を規定する様々なオプションが出てくる。以下ではこの分析手順に沿って,各ステップの詳細と本システムが提供する分析オプションを解説する。また分析のオプションを表1に整理した。

なお,プログラム開発の大部分はJava言語を用い,

サーブレット/JSPおよびJakarta Strutsフレームワーク¹¹⁾を利用してWebアプリケーションとしてシステムを実現した。システム動作時のウィンドウショットを図1,図2に例示する。

3.1 分析対象カリキュラムとシラバスの選択

分析対象としたいカリキュラムを選択し,そのカリキュラム構成内から対象とするシラバス集合を確定するのが最初のステップである。

典型的な分析シナリオは,同じ分野に属する多数のカリキュラムを対象とし,それらを構成するシラバス集合のクラスタリング結果によって,注目するカリキュラムが他との比較においてどのような特徴を持つかを観察することである。また,「授業形態が演習・実験の科目は除く」、「必修の科目のみ」、「科目名が『特論』と

日本の大学ではカリキュラムは大きく「教養」と「専門」に分かれる。本研究では,「専門用語を手がかりとした横断的比較に基づく特徴把握」という方法論により馴染みやすい専門教育のカリキュラムを考察の対象とするが,教養教育カリキュラム¹²⁾に対する適用も興味深い。また,大学によっては学科内に複数の教育コースを設置し,複数のカリキュラムを提供する場合もあるが,本研究ではシラバスデータの整備状況との関係上,1学科が1カリキュラムを提供すると想定している。

いう語を含むものは除く」等の条件で限定されたシラバスを対象にして分析を行うことも考えられる。

これらの選択条件を組み合わせ分析対象シラバス集合を定めて、以下に説明する分析を行い、仮説的なカリキュラムの特徴を見出したら、より絞り込んだ選択条件で対象シラバス集合を限定して分析を繰り返し、カリキュラムの特徴をさらに詳細に検討していく、というのが著者らの想定する本システム利用の流れとなる。

なお本システムが扱うシラバスは、著者らの提案するシラバス XML スキーマ¹³⁾に適合するよう、データ項目を整理したものである。

3.2 対象データ項目の選択とシラバス内容の定量化

分析対象シラバス集合を定めたら、次に各シラバスに出現する専門用語を抽出し、その出現頻度に基づいたシラバスの縮約・定量化された表現を求める。

シラバス XML スキーマ¹³⁾ではシラバスが備える様々なデータ項目を整理している。本システムではその中でテキスト値をとる項目——特に「授業概要」「授業の目的」「授業計画」など、そのシラバスが記述する科目の内容をよく反映していると想定される項目——を選択して、専門用語抽出の対象とすることができる。

専門用語抽出では、専門用語辞書（用語リスト）が必要となる。専門用語辞書の準備については、対象文書コーパスから単語出現の統計情報等に基づき自動生成する様々な方法が提案されており、また既存の辞書を利用することもできる。本システムでは、接続する語の多様性に基づき単名詞・複合名詞の専門性を評価する手法¹⁴⁾を利用した『専門用語自動抽出システム』¹⁵⁾で生成された用語リストと、EDR 専門用語辞書（情報処理）¹⁶⁾のいずれかから選択可能とした。また新たな辞書も自由に追加できる。

各シラバスについて、用語抽出の対象データ項目の文字列を形態素解析システム『茶筌』¹⁷⁾で分かち書きし、用語辞書エントリとのマッチングを判定して、そのシラバスが含む専門用語の種類と頻度を算出する。この情報に基づき、1つのシラバス s_i は

$$s_i = \{(t_k, w(s_i, t_k))\}_{t_k \in T_i} \quad (1)$$

という表現で定量化される。ただし T_i はシラバス s_i が含む専門用語の集合、 $w(s_i, t_k)$ は s_i における

専門用語 t_k の重み付け関数である。本システムでは $w(s_i, t_k)$ として、用語 t_k のシラバス s_i における出現頻度 $tf(t_k, s_i)$ 、および t_k の s_i における TFIDF 値¹⁸⁾ $tfidf(t_k, s_i) = tf(t_k, s_i) \times \log(|S|/df(t_k))$ （ただし $|S|$ は分析対象シラバスの総数、 $df(t_k)$ はそのうち用語 t_k を含むシラバスの総数を表す）を利用可能とした。

なお式 (1) のシラバス定量化表現は、 $t_k \notin T_i$ のとき $w(s_i, t_k) = 0$ とすれば、文書クラスタリングで一般的に用いられるベクトル空間モデル¹⁹⁾（全専門用語 $\cup_i T_i$ が張るベクトル空間、以下用語ベクトル空間と呼ぶ）におけるベクトル表現と等価である。

3.3 シラバス間の類似度計算

シラバス間の類似度を計算する前に、上ステップで得た各シラバスの定量化表現を規格化するか否かを決定する必要がある。シラバスには記述量のばらつきがあり、記述量の多いシラバスほどより多くの専門用語と高い用語重み合計を持つ傾向がある。そこで、すべての科目は同程度の内容を含むはずだと考える場合には、次の規格化操作

$$w(s_i, t_k) \leftarrow \frac{w(s_i, t_k)}{\sum_{t_l \in T_i} w(s_i, t_l)} \quad (2)$$

をすべての用語重みに対して施す。一方、シラバスの記述量が科目内容の量を反映すると考える場合には、規格化は不要である。

規格化の有無を選択したうえで、すべてのシラバスの対 s_i と s_j について類似度 $sim(s_i, s_j)$ を計算する。文書間の類似度の定義にも様々なものが提案されているが、本システムではベクトル空間モデルにおける内積類似度

$$sim_{ip}(s_i, s_j) = \sum_{t_k \in T_i \cap T_j} w(s_i, t_k) \cdot w(s_j, t_k) \quad (3)$$

およびユークリッド距離の逆数

$$sim_{iEuclid}(s_i, s_j) = \left\{ \sum_{t_k \in T_i \cup T_j} ((w(s_i, t_k) - w(s_j, t_k))^2) \right\}^{-1/2} \quad (4)$$

を利用可能とした。なお、シラバスの定量化表現を規格化しているとき、内積類似度はコサイン類似度および Dice 係数に一致する¹⁹⁾。

3.4 シラバスのクラスタリング

得られた類似度を用いてシラバスのクラスタリングを行う。ここで留意したいのは、各シラバスの用語重みによる定量化表現（式 (1)）に立ち戻ってクラスタ-クラスタ間やクラスタ-シラバス間の類似度を何度も計算し直すのは計算資源を余計に消費しがちだという点

シラバスデータはリレーショナルデータベースで管理される。論文など最新の学術情報とは異なり、大学教育におけるシラバスでは、ある程度普及した専門用語しか現れないと予測されるため、既製の専門用語辞書も有効に活用できる。形態素への分かち書きは、たとえば「グラフ」という用語が「ブラグラフ」から抽出されないようするために必要となる。

である。したがってここでは、3.3 節のステップで求めたシラバス間類似度のみを用いてクラスタリングを行える手法が望ましい。そこで本システムでは、文書クラスタリングで主に用いられる階層併合的クラスタリング手法²¹⁾のうち、クラスタ間距離として最短距離法・群平均距離法・最長距離法を利用可能とした。生成するクラスタの個数は、クラスタの併合が起こるクラスタ間類似度の分布を見ながら、樹形図のカットを適切に定めることで設定できる。

なお階層併合的クラスタリングではヒープを用いるアルゴリズム²²⁾を利用することにより、単純なアルゴリズムだと $O(|S|^3)$ にかかる計算量を $O(|S|^2 \log |S|)$ に軽減した。

参考として、Intel Pentium III 800 Mhz のコンピュータ上に設置したシステムで、異なる大きさのシラバス集合に対してクラスタリングを行った際の計算時間を表 2 に示す。

3.5 クラスタへの帰属分布への観察

以上の手続きで、シラバスは少数個のクラスタに分類されるが、カリキュラムの特徴をとらえるうえでは各クラスタの意味内容を素早く理解するためのクラスタ要約が望まれる。クラスタ要約にも様々なバリエーションが考えられているが²³⁾、本システムでは専門用語 t_k のクラスタ C_m の成立への寄与度 ctr_{km} をそのクラスタ内の用語重みの総和

$$ctr_{km} = \sum_{s_i \in C_m} w(s_i, t_k). \quad (5)$$

として定義し、寄与度の高い順に並べた専門用語のリストをクラスタ要約として提供する。

個々のシラバスがどんなクラスタに分類されるかだけでなく、カリキュラム提供機関別、必修/選択の区分、授業形態、開講学年など、シラバスデータの中のカテゴリカルなデータ項目に沿ってシラバスのクラスタへの帰属分布を比較できると、カリキュラムの特徴をより把握しやすくなる。本システムでは、これらのカテゴリカルなデータ項目を組み合わせて分類軸を生成し、その分類軸の各ケースを行、所属クラスタを列とするクロス表を作成する。また、このクロス表にコレスポネンス分析および主成分分析²⁴⁾を適用し、分類軸の各ケースおよび各クラスタを平面上マッピングとして視覚化する。

ただし文書間類似度をコサイン類似度に限定した場合、効率的なアルゴリズムも提案されている²⁰⁾。

ただしコレスポネンス分析では分類軸のケース数が 2 以上、主成分分析では分類軸のケース数がクラスタ数以上でない、平面上マッピングは生成されない。

表 2 異なるサイズのシラバス集合に対するクラスタリング実行時間
Table 2 Clustering execution time for various size of syllabus-sets.

対象シラバス数	60	168	356	687	1333
実行時間 (sec)	1.6	12.5	64.4	257.6	1131.7

なお、コレスポネンス分析および主成分分析によるクラスタ帰属分布の視覚化にはオープンソースの統計ソフト R²⁵⁾ を利用した。

4. 分析の例

最初に 4.1 節で、各分析ステップにおける分析オプションの選択が分析結果にどう影響するかを、実例を交えながら説明する。次に 4.2 節で、本システムを用いてのカリキュラム分析の典型的な流れを例示する。

4.1 分析オプションの影響

分析オプションの影響は、小さなシラバス集合を対象にした方が把握しやすい。そこで本節の分析例では、情報処理学会が提供している『情報系学科のためのコンピュータサイエンス教育カリキュラム J97』²⁶⁾ に含まれている 47 科目(表 3 参照)のシラバスを分析対象とする。専門用語抽出の対象とする項目には「科目名」、「授業概要」、「授業の目的」、「授業計画」、「教科書」を用いる。専門用語辞書には EDR¹⁶⁾ を用いる。また、次のオプション選択を基準オプション設定として固定する：用語重み付け関数として TFIDF 値、シラバス間類似度の算出には式 (2) の規格化を施したうえで式 (3) の内積類似度(コサイン類似度)、クラスタリングには群平均距離法を利用(表 1 では以上のオプション値を太字にしている)。この基準オプション設定で得られるクラスタ要約およびクラスタリング結果(クラスタ数 8)を、それぞれ表 4 および表 5 に示す。クラスタ C_0 はプログラミングやソフトウェア開発、 C_1 は回路設計とアルゴリズム、 C_2 はコンピュータアーキテクチャ、 C_3 はセキュリティと暗号、 C_4 は音声・画像処理、 C_5 は形式言語と文法、 C_6 は符号理論、 C_7 は数値計算、というように、クラスタリングによって直観的にも理解しやすい分類が形成されていることが見てとれる。

以下では、基準オプション設定から用語重み付け関数、シラバス間類似度、クラスタリング手法のそれぞれを変えたときのクラスタリング結果の違いを例示しながら、分析オプション選択の分析結果への影響を説

階層併合的クラスタリングで算出されるクラスタ間類似度に基づき、5 ~ 10 の間でクラスタ分割が最も明確になる値をとった(以下 4.1 節内では比較の容易さのため、クラスタ数 8 に統一)。4.2 節におけるクラスタ数選択も同様の基準による。

表 3 J97²⁶⁾ に含まれる科目コードと科目名

Table 3 Subject codes and names contained in J97²⁶⁾.

L: CS リテラシ科目	
L-1	コンピュータサイエンス序説
L-2	プログラミング入門
M: 学部科目 (情報数学系)	
M-1	情報代数と符号理論
M-2	離散数学
M-3	計算論
M-4	確率論と情報理論
M-5	数理計画法
M-6	数理論理学
U: 学部科目	
U-1	論理回路
U-2	形式言語とオートマトン
U-3	データ構造とアルゴリズム
U-4	コンピュータアーキテクチャ
U-5	プログラミング言語論
U-6	論理設計
U-7	オペレーティングシステム
U-8	コンパイラ
U-9	デジタル通信
U-10	データベース
U-11	人工知能
U-12	情報ネットワーク
U-13	ソフトウェア工学
U-14	数値計算
U-15	集積回路工学
U-16	信号処理
U-17	画像情報処理
U-18	パターン認識
U-19	ヒューマン—コンピュータ・インタラクション
U-20	コンピュータグラフィックス
U-21	自然言語処理
A: 大学院科目 (学部科目の特論)	
A-1	論理回路特論
A-2	コンピュータアーキテクチャ特論
A-3	オペレーティングシステム特論
A-4	データベース特論
A-5	画像情報処理特論
A-6	コンピュータグラフィックス特論
A-7	信号処理特論
A-8	ソフトウェア工学特論
G: 大学院科目	
G-1	ハードウェアアルゴリズム論
G-2	論理設計と CAD
G-3	プログラム意味論
G-4	ネットワークアーキテクチャ
G-5	分散処理システム
G-6	情報セキュリティ
G-7	関数プログラミング
G-8	オブジェクト指向プログラミング
G-9	音声情報処理
G-10	アルゴリズム論

表 4 基準オプション設定でのクラスタ要約

Table 4 Obtained cluster summaries with standard option set.

C_0	オブジェクト, オブジェクト指向, データベース, ...
C_1	回路, 論理, 設計, アルゴリズム, 計画, 集積回路, ...
C_2	コンピュータ, 通信, 分散, プロセス, システム, ...
C_3	確率, 暗号, 安全性, 情報セキュリティ, 鍵, セキュ...
C_4	画像, 識別, フィルタ, 変換, 信号, 音声, 次元, ア...
C_5	文法, 解析, 構文解析, 文脈, オートマトン, 構文, ...
C_6	体, 巡回, 巡回符号, 符号理論, ガロア体, 線形符号, ...
C_7	数値, 補間, 推定, 桁落ち, 桁, 離散化誤差, 非線形...

表 5 基準オプション設定でのクラスタリング結果

Table 5 Result of clustering with standard option set.

C_0	L-2, M-3, U-5, U-10, U-13, A-4, A-8, G-3, G-7, G-8
C_1	M-2, M-5, M-6, U-1, U-3, U-6, U-11, U-15, A-1, G-1, G-2, G-10
C_2	L-1, U-4, U-7, U-9, U-12, U-19, A-2, A-3, G-4, G-5
C_3	M-4, G-6
C_4	U-16, U-17, U-18, U-20, A-5, A-6, A-7, G-9
C_5	U-2, U-8, U-21
C_6	M-1
C_7	U-14

表 6 用語重み付け関数に tf を用いた結果のクラスタ要約

Table 6 Obtained cluster summaries with tf for term weighting.

C_0	プログラム, オブジェクト, 言語, オブジェクト指...
C_1	計算, アルゴリズム, 問題, 関数, モデル, 項目, 計...
C_2	コンピュータ, システム, 分散, プロセス, インタ...
C_3	論理, 回路, 設計, 関数, 項目, 集積回路, 手法, 論...
C_4	情報, 理論, 確率, 暗号, 技術, 安全性, 項目, 知識, ...
C_5	画像, 変換, 処理, 信号, フィルタ, 表現, 次元, ディ...
C_6	言語, 解析, 文法, 文脈, オートマトン, 構文, 問題, ...
C_7	ネットワーク, 制御, 通信, 技術, コンピュータネッ...

表 7 用語重み付け関数に tf を用いた時のクラスタリング結果

Table 7 Result of clustering with tf for term weighting.

C_0	L-2, U-5, U-10, U-13, A-4, G-3, G-8
C_1	M-3, M-5, U-3, U-14, U-18, G-1, G-7, G-9, G-10
C_2	L-1, U-4, U-7, U-19, A-2, A-3, A-8, G-5
C_3	M-2, M-6, U-1, U-6, U-11, U-15, A-1, G-2
C_4	M-1, M-4, G-6
C_5	U-16, U-17, U-20, A-5, A-6, A-7
C_6	U-2, U-8, U-21
C_7	U-9, U-12, G-4

グ結果を, それぞれ表 6 および表 7 に示す. この結果は, 表 4, 表 5 の基準オプション設定での結果と比べても, 直観的に同程度まで納得できるものである.

ただし, 学問分野などの事情で既存の専門用語辞書を手軽に利用できない場合は状況が異なる. その場合、『専門用語自動抽出システム』¹⁵⁾ のように対象文書コー

明する.

用語重み付け関数

用語重み付け関数 w として用語の出現頻度 tf を用いたとき, 得られるクラスタ要約およびクラスタリン

パスから専門用語リストを自動生成²⁷⁾するツールを用いることになるが、これらのツールでは「こと」「項目」「目的」など高頻度で出現する非専門用語も抽出されることが多い。このような場合には、用語重み付け関数に *tfidf* を用いると、非専門用語の影響をある程度抑えることができる。

シラバス間類似度

用いるシラバス間類似度 *sim* の選択が分析結果を大きく左右するのは、式 (2) の規格化を行わない場合である。規格化なしで式 (3) の内積類似度を用いたときのクラスタリング結果を表 8 に示す。また、規格化なしで式 (4) のユークリッド距離の逆数類似度を用いたときのクラスタリング結果を表 9 に示す(ともにクラスタ要約は省略)。

表 8 と表 9 を比較すると、クラスタリング結果が大きく違うことが分かる。この違いを理解するためには、それぞれ式 (3), (4) の定義より、内積類似度では共有する用語の多さが類似度を高めるのに対して、ユークリッド距離の逆数では共有しない用語の少なさが類似度を高める、という点に注意すればよい。これより規格化なしの内積類似度では、含む用語の絶対数が少ない—用語ベクトル空間の原点に近い—シラバスは、他のシラバスとの類似度が低くなるため、それ自身のみを含む孤立したクラスタを形成しがちとなる。他方、規格化なしのユークリッド距離の逆数では、含む用語

の多いシラバスほど他のシラバスとの類似度が低くなり、孤立したクラスタを形成しがちとなる。以上のことは、

$$norm(s_i) = \left\{ \sum_{t_k \in T_i} w(s_i, t_k)^2 \right\}^{1/2} \quad (6)$$

で求められる各シラバスのユークリッドノルムと照らし合わせたとき、表 8 で孤立または小さなクラスタ ($C_2, C_4 \sim C_7$) に含まれるシラバスの多くが小さなノルムを持ち、逆に表 9 で孤立したクラスタ ($C_1 \sim C_7$) に含まれるシラバスが大きなノルムを持つこと(表 10 参照) から確認できる。

このことから、

- 「抽出される用語が少ないシラバスは、分野の平均から離れたそのカリキュラム独自の科目である」と考え、その存在を強調したいときには内積類似度、
- 逆に「多くの用語を含むシラバスはカリキュラムの独自性を代表する科目である」と考え、その存在を強調したいときにはユークリッド距離の逆数類似度、

といった使い分けも考えられる。

なお、式 (2) の規格化を用いた場合にも、ユークリッド距離の逆数を用いた場合は、内積類似度を用いた場合に比べて、「シラバスの大多数を含む大きなクラスタが 1 つ、残りは小さなクラスタ」という結果が得られやすい傾向が見られたことを付記しておく。

クラスタリング手法

最短距離法を用いたクラスタリングで得られるクラスタリング結果を表 11 に示す。また、最長距離法を

表 8 規格化なし、内積類似度でのクラスタリング結果
Table 8 Result of clustering with non-normalization, inner-product-distance for similarity.

C_0	L-2, M-2, M-6, U-1, U-2, U-3, U-5, U-6, U-8, U-10, U-11, U-15, U-21, A-1, A-4, G-1, G-2, G-8, G-10
C_1	L-1, U-4, U-7, U-9, U-12, U-19, A-2, A-3, G-4, G-5
C_2	M-4, G-6
C_3	U-14, U-16, U-17, U-18, U-20, A-5, A-6, A-7, G-9
C_4	M-3, G-7
C_5	M-1
C_6	U-13, A-8, G-3
C_7	M-5

表 9 規格化なし、ユークリッド距離の逆数類似度でのクラスタリング結果

Table 9 Result of clustering with non-normalization, inverse-Euclid-distance for similarity.

C_0	$C_1 \sim C_7$ 以外の 45 科目	C_1	U-16
C_2	A-7	C_3	U-11
C_4	U-8	C_5	M-6
C_6	U-14	C_7	U-4

表 10 各シラバスの用語ベクトル空間でのユークリッドノルム
Table 10 Euclidean norm of each syllabus in term-vector-space.

L-1	50.8	L-2	78.4	M-1	81.0
M-2	46.4	M-3	54.4	M-4	88.6
M-5	60.9	M-6	129.0	U-1	95.8
U-2	84.6	U-3	54.8	U-4	115.9
U-5	46.5	U-6	73.0	U-7	65.2
U-8	141.5	U-9	92.0	U-10	69.6
U-11	109.5	U-12	65.4	U-13	34.5
U-14	106.9	U-15	67.3	U-16	110.6
U-17	100.4	U-18	99.2	U-19	82.6
U-20	59.8	U-21	87.4	A-1	69.2
A-2	52.8	A-3	60.3	A-4	67.3
A-5	64.5	A-6	30.7	A-7	116.2
A-8	60.1	G-1	64.5	G-2	70.9
G-3	47.9	G-4	71.6	G-5	67.3
G-6	83.1	G-7	42.8	G-8	73.3
G-9	96.0	G-10	31.3		

用いたクラスタリングで得られるクラスタリング結果を表 12 に示す（ともにクラスタ要約は省略）。

表 11 と表 5 の対比からも分かるように、最短距離法はシラバスの大多数を含む大きなクラスタを形成する傾向がある。これは近接するシラバスが同一のクラスタに次々に取り込まれていく結果、用語ベクトル空間内で鎖状に長く伸び、内部にあまり類似しないシラバス対も含むクラスタが生成されやすい、“chain effect” と呼ばれる性質のためである。逆に最長距離法は、表 12 にも見られるように、均一な大きさのクラスタを生成する傾向がある。これは、内部に多くのシラバスを含むクラスタほど、新たなシラバスとの距離は大きくなる可能性が高くなり、その結果として極端に大きなクラスタの形成が抑制されるためである。

このことから、

- 他と顕著に違う少数の科目を発見したいときには最短距離法、
- 均一な大きさの分類が自然に期待されるシラバス集合で、分類間のバランスをカリキュラムどうしで比較したいときには最長距離法、

といった使い分けも考えられる。

なお、本システムで利用できる 3 つを含む種々の階層的クラスタリング手法の妥当性について、より定量的な議論も行われている²⁰⁾。ただし 2 章でも言及したように、クラスタリングは個々のデータ内容のみに基づいてデータセットに分類構造を定める分析手法である。したがって、外的に定められた分類クラスの再

表 11 クラスタリングに最短距離法を用いた時のクラスタリング結果

Table 11 Result of clustering with single-link method for clustering.

C ₀	C ₁ ~ C ₇ 以外の 30 科目	C ₁	L-1
C ₂	M-4, U-16, U-17, U-18, U-20, A-5, A-6, A-7, G-9	C ₃	G-6
C ₄	U-2, U-8, U-21	C ₅	M-1
C ₆	L-2	C ₇	U-14

表 12 クラスタリングに最長距離法を用いた時のクラスタリング結果

Table 12 Result of clustering with complete-link method for clustering.

C ₀	L-2, U-5, U-10, U-13, A-4, G-3, G-8
C ₁	M-3, U-3, U-14, G-1, G-7, G-10
C ₂	L-1, M-5, U-11, U-18, A-5
C ₃	M-2, M-6, U-1, U-6, U-15, A-1, G-2
C ₄	U-4, U-7, U-12, U-19, A-2, A-3, A-8, G-4, G-5
C ₅	M-1, M-4, U-9, G-6
C ₆	U-16, U-17, U-20, A-6, A-7, G-9
C ₇	U-2, U-8, U-21

現率等が手法の絶対的な評価基準とはならないこと、分析の文脈に合わせてそのつど、適切な手法を検討する余地があることは強調しておきたい。

4.2 典型的な分析の流れ

3.1 節でも述べたように、本システムを用いた典型的な分析は、最初に複数のカリキュラムを対象にしてその中での注目するカリキュラムの相対的な位置付けを把握し、その結果を元に対象シラバス集合を限定、かつ観察の分類軸を詳細にしていきながら分析を繰り返すことで、カリキュラムの特徴をより詳しく把握していく、という流れに従う。以下では、この流れに従った分析例を 1 つ示す。なお分析オプションは、一貫して前の 4.1 節の基準オプション設定を用いている。

まず最初に全国の国立大学の中から、Web 上にシラバスが公開されており学科名が「情報」と「工学」を含む 9 大学 10 学科をランダムに選び、それらの学科が提供しているカリキュラムのシラバス集合（総数 687 科目、2002 年度版）を分析対象とした。

クラスタリング（クラスタ数 7）で得られたクラスタ要約を表 13 に示す。また、学科を分類軸にしてのクロス表を表 14 に示す（“d#” が各学科を表す）。さらにコレスポンデンス分析を用いてこのクロス表を平面上マッピングした結果を図 3 に示す（図中、“C.#” が各クラスタを表す）。これらの結果より、

表 13 10 学科を対象とした分析でのクラスタ要約
Table 13 Cluster summaries analysing curriculums of 10 departments.

C ₀	プログラミング, 言語, 情報, 技術, アルゴリズム, ...
C ₁	微分, 微分方程式, 展開, 積分, 確率, 関数, 数値, ...
C ₂	回路, 論理, 電気, 電子, 原子, エネルギー, 光, 半...
C ₃	信号, 画像, 通信, パターン, 符号化, デジタル, ...
C ₄	暗号, 群, セキュリティ, 鍵, 同型, 多項式, 体, 情...
C ₅	アーキテクチャ, 計算機, 管理, オペレーティング...
C ₆	生体, 知覚, 神経, インタフェース, コミュニケーション...

表 14 10 学科を対象とした分析での学科—クラスタのクロス表
Table 14 Department-cluster cross table analysing curriculums of 10 departments.

	C ₀	C ₁	C ₂	C ₃	C ₄	C ₅	C ₆	計
d0	17	24	9	8	1	3	2	64
d1	19	4	8	5	1	4	1	42
d2	30	16	25	9	2	3	0	85
d3	21	8	16	11	0	3	1	60
d4	27	25	8	8	2	3	0	73
d5	15	21	28	7	2	3	3	79
d6	51	28	10	4	3	2	1	99
d7	21	7	61	7	0	2	2	100
d8	20	11	13	5	0	2	2	53
d9	15	5	6	3	0	2	1	32
計	236	149	184	67	11	27	13	687

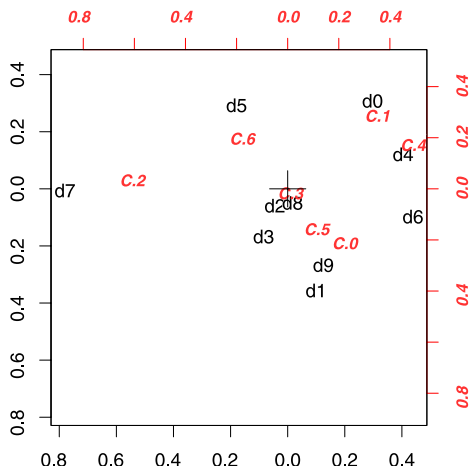


図 3 10 学科を対象にした分析での、学科—クラスタのコレスポネンス分析

Fig.3 Correspondence analysis for 10 departments, 7 clusters.

- C_1 は連続変数を扱う数学系の科目を多く含むクラスタであり、学科 d0 の提供するカリキュラムはその方面に力を入れていること、
 - C_2 は電子回路や電子材料系の科目を多く含むクラスタであり、学科 d7 は特にその方面に力を入れていること、
 - C_0, C_3, C_5 はそれぞれソフトウェア、信号処理、ハードウェア系の科目を多く含むクラスタであり、学科 d1, d2, d3, d8, d9 はいずれも、これらの科目を平均的に備えた、情報工学系として平均的なカリキュラムと考えられること、
 - C_4 は暗号・セキュリティの科目、 C_6 は生体情報処理やヒューマンインタフェースの科目からなるクラスタであり、これらの科目を備えていない学科もあること、
- などが読み取れる。

次にこれらの学科の中から、学科名が同じであり、提供する科目数も近い 2 つの学科 d0 と d3 を取り出し、さらに詳細な分析を加えた。まず、クラスターリング(クラスタ数 7) で得られたクラスタ要約を表 15 に、上と同じく学科を分類軸にしてのクロス表を表 16 に示す。これらより、

- C_1 は数学系の科目を多く含むクラスタであり、学科 d0 は d3 よりもこの方面に力を入れていること(上の分析でも確認したとおり)、
 - C_4 は電磁波に関係する科目からなるクラスタであり、学科 d3 はその方面の科目を特異的に揃えていること、
- などが読み取れる。

表 15 2 学科を対象とした分析でのクラスタ要約

Table 15 Cluster summaries analysing curriculums of 2 departments.

C_0	回路、技術、設計、会話、情報、システム、通信、信...
C_1	展開、プログラミング、微分、積分、問題、微分方程...
C_2	アルゴリズム、コンピュータグラフィックス、画像、...
C_3	オペレーティングシステム、管理、装置、方式、命令...
C_4	電波、無線、アンテナ、磁気、利用、電界、静電界、...
C_5	意味論、仕様記述、形式的仕様、形式的仕様記述、仕...
C_6	率、測定、干渉、分光計、放電、オシロスコープ、温...

表 16 2 学科を対象とした分析での学科—クラスタのクロス表

Table 16 Department-cluster cross table analysing curriculums of 2 departments.

	C_0	C_1	C_2	C_3	C_4	C_5	C_6	計
d0	21	30	6	5	0	1	1	64
d3	33	13	6	3	5	0	0	60
計	54	43	12	8	5	1	1	124

表 17 表 16 を必修/選択区分 (R/S) で詳細化したクロス表

Table 17 Refinement of cross table 16 with required/selective category.

	C_0	C_1	C_2	C_3	C_4	C_5	C_6	計
d0R	3	5	0	1	0	0	0	9
d0A	11	19	4	4	0	1	1	40
d0S	7	6	2	0	0	0	0	15
d3R	12	3	0	1	1	0	0	17
d3S	21	10	6	2	4	0	0	43
計	54	43	12	8	5	1	1	124

さらに同じクラスターリング結果を、学科と必修/選択区分を組み合わせた分類軸に沿って集計したクロス表を表 17 に、コレスポネンス分析を用いてこのクロス表を平面上マッピングした結果を図 4 に示す(“d#R” が必修科目、“d#A” は「指定選択科目」、 “d#S” が選択科目を表す)。これらより、

- 学科 d0 に特徴的であった数学系の科目(クラスタ C_1) は、その多くが「指定選択科目」という区分に属すること、
 - 学科 d3 に特徴的であった電磁波に関係する科目(クラスタ C_4) は、その多くが選択科目であること、
 - そのほか、クラスタ C_5 や C_6 を形成する d0 の独特な科目も、「指定選択科目」に属すること、
- などが読み取れる。

このほか、開講学年や授業形態といったカテゴリカルなデータ項目を組み合わせた分類軸に沿った分析を行うことで、カリキュラムの特徴をさらに詳細に把握することができる。

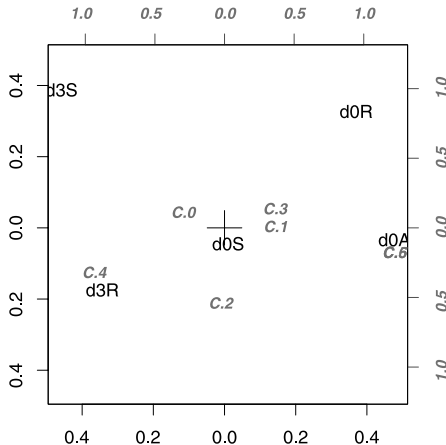


図4 2学科対象での、学科×必修/選択区分—クラスタの相対分析

Fig.4 Correspondence analysis for 5 cases of department-requirement combination, and 7 clusters.

5. まとめと今後の課題

本研究では、カリキュラムを構成する授業科目のシラバスをその記述内容の類似性に基づきクラスタリングし、シラバスのクラスタへの帰属分布を様々な分類軸に沿って比較することで、カリキュラムの特徴把握を促進するシステムを構築した。本システムが提供する分析の詳細を解説し、本システムを用いてのカリキュラム分析の例を示した。

3章で示したように、本システムのカリキュラム分析には多数のオプションがある。そして4章で示したように、分析オプションの選択によって分析結果の様相はときに大きく変わってくる。しかし、このような分析結果の不定性・多様な解釈の可能性は、複雑な情報源からの知識発見にはつきものである。むしろカリキュラム設計者や評価者の視点を反映した幅広いオプションを試しては分析結果を観察するという繰返し、多角的な視点からのカリキュラム理解には必要であろう。なお、「標準的なオプションの組合せ」については今後システムの利用実績を重ね、分析の観点ごとに使用者の意見を蓄積していきたい。

知識発見という用途上、システムの有効性を定量的に評価することは困難である。しかし、カリキュラム設計に経験のある専門教育者にシステムを試用してもらったところ、「単にシラバスを読み通していくのと比べて、カリキュラムの全体像がより把握しやすい」、「あるカリキュラムに特徴的な科目を素早く発見できた」、「どの方向の科目を強めればカリキュラムの個性が出せるか、示唆を与えてくれる」といった肯定的な

感想を得ることができた。

他方、本システムに改善を望む意見として、大量のシラバスを分析対象としたときのクラスタリングの高速化があげられた。実際、4.2節で例示したような繰返し分析をストレスなく行うために、対象シラバス集合の選択から分析結果取得までのターンアラウンドの短縮は重要である。そこで現在、より高速な階層分割的クラスタリング手法²⁰⁾などの導入を検討している。

本システムでは、シラバスのクラスタリングを記述内容の類似性に基づき行う。しかしシラバス間には記述の類似性という関係のほかに、「あらかじめ履修しておくべき科目」、「同時に履修すべき科目」といった履修上の関係もあり、カリキュラムの設計や分析では各科目の内容記述とならんで重要な情報であると考えられる。そこで今後は、2章でも言及した文書間リンクに基づく分析手法と、本研究の分析手法とを組み合わせたい。

謝辞 本研究を遂行するにあたりご協力いただいた大学評価・学位授与機構「大学評価情報の構造解析と評価プロセスへの応用の研究会」参加者の皆様に謝意を表します。

参考文献

- 1) 内山弘美：環境冠学科の設置メカニズム—国立大学工学系学部を事例として、高等教育ジャーナル—高等教育と生涯学習、Vol.8, pp.1-15 (2000).
- 2) 大場善次郎：東大工学教育プロジェクト「工学知の構造化と可視化」、カレッジマネジメント、No.119, pp.34-38 (2003).
- 3) Miyazaki, K., Ida, M., Yoshikane, F., Nozawa, T. and Kita, H.: On Development of a Course Classification System using Syllabus Data, *Proc. ICOME2003*, pp.68-69 (2003).
- 4) 高須淳宏, 相原健郎：テキスト分類における訓練データと性能の実験的考察, *NII Journal*, Vol.6, pp.1-8 (2003).
- 5) Willett, P.: Recent trends in hierarchical document clustering: A critical review, *Information Processing and Management*, Vol.24, No.5, pp.577-597 (1988).
- 6) Cutting, D.R., Pedersen, J.O., Karger, D. and Tukey, J.W.: Scatter/Gather: A Cluster-based Approach to Browsing Large Document Collections, *Proc. 15th Annual Int'l ACM SIGIR Conf. on R&D in IR*, pp.318-329 (1992).
- 7) Hearst, M.A. and Pedersen, J.O.: Reexamining the cluster hypothesis: Scatter/gather on retrieval results, *Proc. SIGIR-96, 19th ACM Int'l Conf. on R&D in IR*, pp.76-84 (1996).
- 8) Small, H. and Sweeney, E.: Clustering the Sci-

- ence Citation Index Using Co-Citations I: A Comparison of Methods, *Scientometrics*, Vol.7, pp.391-409 (1985).
- 9) Botafogo, R.A.: Cluster analysis for hypertext systems, *Proc. 16th Annual Int'l ACM SIGIR Conf. on R&D in IR*, pp.116-125 (1993).
- 10) Weiss, R., Vélez, B., Sheldon, M., Nemprempre, C., Szilagyi, P., Duda, A. and Gifford, D.K.: HyPursuit: A hierarchical network search engine that exploits content-link hypertext clustering, *Proc. 7th ACM Conf. on Hypertext*, pp.180-193 (1996).
- 11) Apache Struts Web Application Framework. <http://jakarta.apache.org/struts/index.html>
- 12) 大学評価・学位授与機構：平成13年度着手分全学テーマ別評価「教養教育」評価報告書(2002). <http://www.niad.ac.jp/>
- 13) 井田正明, 宮崎和光, 芳鐘冬樹, 喜多一：シラバスXMLデータベースシステム構築に関する考察, 情報処理学会第65回全国大会 2A-6, Vol.4, pp.4-247-4-248 (2003).
- 14) 湯本紘彰, 森辰則, 中川裕志：出現頻度と接続頻度に基づく専門用語抽出, 情報処理学会第145回自然言語処理研究会, pp.111-118 (2001).
- 15) 東京大学中川研究室・横浜国立大学森研究室：専門用語自動抽出システム。
- 16) (株)日本電子化辞書研究所：EDR電子化辞書専門用語辞書(情報処理)。
- 17) 松本裕治, 北内啓, 山下達雄, 平野善隆, 松田寛, 高岡一馬, 浅原正幸：形態素解析システム『茶釜』version 2.2.1 使用説明書, NAIST Technical Report, NAIST-IS-TR99012, 奈良先端科学技術大学院大学(1999).
- 18) Salton, G.: Developments in automatic text retrieval, *Science*, Vol.253, pp.974-980 (1991).
- 19) Salton, G.: *Automatic Text Processing: The transformation, Analysis and Retrieval of Information by Computer*, Addison-Wesley (1989).
- 20) Zhao, Y. and Karypis, G.: Evaluation of Hierarchical Clustering Algorithm for Document Datasets, *Proc. 11th Int'l Conf. on Information and Knowledge Management*, pp.515-524 (2002).
- 21) 宮本定明：クラスター分析入門：ファジィクラスタリングの理論と応用, 森北出版(1999).
- 22) Kurita, T.: An efficient agglomerative clustering algorithm using a heap, *Pattern Recognition*, Vol.24, No.3, pp.205-209 (1991).
- 23) 森辰則：検索結果表示向け文書要約における情報利得比に基づく語の重要度計算, 自然言語処理, Vol.9, No.4, pp.3-32 (2002).
- 24) 大隅昇：記述的多変量解析法, 日科技連出版社(1994).

- 25) 中澤 港：Rによる統計解析の基礎, ピアソン・エデュケーション(2003).
- 26) (社)情報処理学会：大学の理工系学部情報系学科のためのコンピュータサイエンス教育カリキュラム J97(第1.1版)(1999年9月)
- 27) 芳鐘冬樹, 井田正明, 宮崎和光, 野澤孝之, 喜多一：シラバスからの専門用語抽出手法の検討, 情報処理学会第66回全国大会, Vol.4, pp.4-375-4-376 (2004).

(平成16年7月7日受付)

(平成16年11月1日採録)



野澤 孝之

昭和50年生。平成14年東京工業大学大学院総合理工学研究科知能システム科学専攻博士課程修了。平成14年から15年にかけて京都産業大学客員研究員。平成15年より大学評価・学位授与機構評価研究部助手。非線型力学, 複雑系の理論とシミュレーション, ネットワーク構造分析に関する研究に従事。博士(理学)。電子情報通信学会会員。



井田 正明(正会員)

昭和39年生。平成2年京都大学大学院工学研究科精密工学専攻修士課程修了。同年京都大学工学部助手。平成14年より大学評価・学位授与機構評価研究部助教授。システム工学, 評価と意思決定, 数理計画法, データベースに関する研究に従事。博士(工学)。計測自動制御学会会員。



芳鐘 冬樹(正会員)

昭和44年生。平成14年東京大学大学院教育学研究科総合教育科学専攻博士課程単位取得退学。同年大学評価・学位授与機構評価研究部助手。研究コミュニティの構造分析, 専門用語の自動抽出手法に関する研究に従事。修士(教育学)。日本図書館情報学会会員。



宮崎 和光 (正会員)

昭和 42 年生。平成 8 年東京工業大学大学院総合理工学研究科知能科学専攻博士後期課程修了。同年同大学大学院総合理工学研究科助手。平成 10 年から 11 年にかけて同大学大学院総合理工学研究科リサーチアソシエイト。平成 11 年学位授与機構審査研究部助教授を経て平成 12 年より大学評価・学位授与機構学位審査研究部助教授。機械学習，特に強化学習に関する研究に従事。博士(工学)。人工知能学会，計測自動制御学会，日本機械学会，日本高等教育学会等の会員。



喜多 一

昭和 34 年生。昭和 57 年京都大学工学部卒業。昭和 62 年京都大学大学院工学研究科博士後期課程研究指導認定退学。昭和 62 年京都大学工学部助手。平成 9 年東京工業大学大学院総合理工学研究科助教授。平成 12 年大学評価・学位授与機構教授を経て平成 15 年京都大学学術情報メディアセンター教授。社会システム，ニューラルネットワーク，進化的計算，エージェントシミュレーションの研究に従事。工学博士。電気学会，電子情報通信学会，計測自動制御学会，システム制御情報学会等の会員。